An $\mathcal{O}(nL)$ Infeasible-Interior-Point Algorithm for Linear Programming

Yuagang Yang*and Makoto Yamashita†

July 1, 2015

Abstract

In this paper, we propose an arc-search infeasible-interior-point algorithm. We show that this algorithm is polynomial and the polynomial bound is $\mathcal{O}(nL)$ which is at least as good as the best existing bound for infeasible-interior-point algorithms for linear programming.

Keywords: polynomial algorithm, infeasible-interior-point method, linear programming.

*US NRC, Office of Research, 21 Church Street, Rockville, 20850. Email: yaguang.yang@verizon.net.

†Department of Mathematical and Computing Sciences, Tokyo Institute of Technology. Email: Makoto.Yamashita@is.titech.ac.jp.

1 Introduction

Since Klee and Minty [1] showed that a simplex method for linear programming is not a polynomial algorithm, polynomial complexity bound has become a popular metric to measure the efficiency of optimization algorithms. Searching for polynomial algorithms for linear programming was a major research area of optimization between 1980's and 1990's after Khachiyan [2] announced the first polynomial algorithm for linear programming. Although Khachiyan's algorithm was shown to be much less efficient in practice than the simplex method [3], Karmarkar's interior-point method [4] demonstrated the possibility of existence of efficient polynomial algorithms. For feasible starting point, people quickly established polynomial bounds for various interior-point algorithms [5, 6, 7, 8, 9, 10]. The lowest bound of these algorithms is $\mathcal{O}(\sqrt{n}L)$ which has not been improved for more than two decades.

To have an efficient implementation for interior-point algorithms, Mehrotra [11] and Lustig et. al. [12] realized that higher-order method and infeasible starting point are two necessary improvements. However, algorithms with either one of these features had poorer complexity bounds than $\mathcal{O}(\sqrt{n}L)$. Monteiro, Adler, and Resende [13] showed that a higher-order algorithm starting from a feasible point has the polynomial bound $\mathcal{O}(nL)$. For infeasible-interior-point method, Zhang [14], Mizuno [15], and Miao [16] established polynomiality for several different algorithms (none of them is a higher-order algorithm). The best complexity bound $\mathcal{O}(nL)$ for infeasible interior-point methods has not been changed since the eary of 1990's.

Recently, Yang [17, 18] showed that for a higher-order interior-point method starting from a feasible point, the polynomial bound can be improved to $\mathcal{O}(\sqrt{n}L)$ by using an arc-search method. Very recently, Yang et. al. [19] used the same idea and proposed a polynomial arc-search infeasible-interior-point algorihtm with a complexity bound of $\mathcal{O}(n^{\frac{5}{4}}L)$. In this paper, we show that for higher-order infeasible-interior-point method using arc-search, the polynomial bound can be improved to $\mathcal{O}(nL)$, which is a bound at least as good as the best bound of existing infeasible-interior-point algorithms.

The remainder of the paper is organized as follows. Section 2 describes the problem. Section 3 provides an infeasible-predictor-corrector algorithm. Section 4 proves its polynomiality. Section 5 summarizes the conclusions.

2 Problem Descriptions

The standard form of linear programming in this paper is given as follows:

$$\min \ \mathbf{c}^{\mathrm{T}}\mathbf{x}, \quad \text{subject to} \ \ \mathbf{A}\mathbf{x} = \mathbf{b}, \ \ \mathbf{x} \geq 0, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$ are given, and $\mathbf{x} \in \mathbb{R}^n$ is the vector to be optimized. Associated with the linear programming is the dual programming that is also presented in the standard form:

$$\max \ \mathbf{b}^{\mathrm{T}}\mathbf{y}, \quad \text{subject to} \ \ \mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{s} = \mathbf{c}, \ \ \mathbf{s} \geq 0, \tag{2}$$

where dual variable vector $\mathbf{y} \in \mathbb{R}^m$, and dual slack vector $\mathbf{s} \in \mathbb{R}^n$. We use $\mathcal{S}$ to denote the set of all the optimal solutions $(\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*)$ of (1) and (2). It is well known that $\mathbf{x} \in \mathbb{R}^n$ is an optimal solution of (1) if and only if $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{s}$ satisfy the following KKT conditions

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{3a}$$
$$\mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{s} = \mathbf{c}, \tag{3b}$$
$$(\mathbf{x}, \mathbf{s}) \geq 0, \tag{3c}$$
$$x_i s_i = 0, \quad i = 1, \ldots, n. \tag{3d}$$

To simplify the notation, we will denote Hadamard (element-wise) product of two vectors $\mathbf{x}$ and $\mathbf{s}$ by $\mathbf{x} \circ \mathbf{s}$, the element-wise division of the two vectors by $\mathbf{s}^{-1} \circ \mathbf{x}$, or $\mathbf{x} \circ \mathbf{s}^{-1}$, or $\frac{\mathbf{x}}{\mathbf{s}}$ if $\min |s_i| > 0$, the Euclidean norm of x by $\|\mathbf{x}\|$, the infinite norm of $\mathbf{x}$ by $\|\mathbf{x}\|_\infty$, the identity matrix of any dimension by $\mathbf{I}$, the vector of all ones with appropriate dimension by $\mathbf{e}$, the block column vectors, for example, $[\mathbf{x}^{\mathrm{T}}, \mathbf{s}^{\mathrm{T}}]^{\mathrm{T}}$ by $(\mathbf{x}, \mathbf{s})$. For $\mathbf{x} \in \mathbb{R}^n$, we will denote a related diagonal matrix by $\mathbf{X} \in \mathbb{R}^{n \times n}$ whose diagonal elements are the components of the vector $\mathbf{x}$. Finally, we define an initial vector point of a sequence by $\mathbf{x}^0$, an initial scalar point of a sequence by μ_0, the vector point after the kth iteration by $\mathbf{x}^k$, the scalar point after the kth iteration by μ_k. Let

$$\mathbf{r}_b^k = \mathbf{A}\mathbf{x}^k - \mathbf{b}, \tag{4a}$$
$$\mathbf{r}_c^k = \mathbf{A}^{\mathrm{T}}\mathbf{y}^k + \mathbf{s}^k - \mathbf{c}. \tag{4b}$$

Given a strictly positive current point $(\mathbf{x}^k, \mathbf{s}^k) > 0$, the infeasible-predictor-corrector algorithm is to find the solution of (1) approximately along a curve $\mathcal{C}(t)$ defined by the following system

$$\mathbf{A}\mathbf{x}(t) - \mathbf{b} = t\mathbf{r}_b^k \tag{5a}$$
$$\mathbf{A}^{\mathrm{T}}\mathbf{y}(t) + \mathbf{s}(t) - c = t\mathbf{r}_c^k \tag{5b}$$
$$\mathbf{x}(t) \circ \mathbf{s}(t) = t\mathbf{x}^k \circ \mathbf{s}^k \tag{5c}$$
$$(\mathbf{x}(t), \mathbf{s}(t)) > 0, \tag{5d}$$

where $t \in (0, 1]$. As $t \to 0$, $(\mathbf{x}(t), \mathbf{y}(t), \mathbf{s}(t))$ appraches the solution of (1). Since $\mathcal{C}(t)$ is not easy to obtain, we will use an ellipse $\mathcal{E}$ [20] in the $2n + m$ dimensional space to approximate the curve defined by (5), where $\mathcal{E}$ is given by

$$\mathcal{E} = \{(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) : (\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) = \vec{\mathbf{a}} \cos(\alpha) + \vec{\mathbf{b}} \sin(\alpha) + \vec{\mathbf{c}}\}, \tag{6}$$

$\vec{\mathbf{a}} \in \mathbb{R}^{2n+m}$ and $\vec{\mathbf{b}} \in \mathbb{R}^{2n+m}$ are the axes of the ellipse, and $\vec{\mathbf{c}} \in \mathbb{R}^{2n+m}$ is the center of the ellipse. Taking the derivatives of (5) gives

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{\mathrm{T}} & \mathbf{I} \\ \mathbf{S}^k & \mathbf{0} & \mathbf{X}^k \end{bmatrix} \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \\ \dot{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_b^k \\ \mathbf{r}_c^k \\ \mathbf{x}^k \circ \mathbf{s}^k \end{bmatrix}, \tag{7}$$

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{\mathrm{T}} & \mathbf{I} \\ \mathbf{S}^k & \mathbf{0} & \mathbf{X}^k \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{x}} \\ \ddot{\mathbf{y}} \\ \ddot{\mathbf{s}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ -2\dot{\mathbf{x}} \circ \dot{\mathbf{s}} \end{bmatrix}. \tag{8}$$

We require the ellipse to pass the same point $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k)$ on $\mathcal{C}(t)$ and to have the same derivatives given by (7) and (8). The ellipse is given in [17, 18] as

Theorem 2.1 *Let* $(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha))$ *be an arc defined by (6) passing through a point* $(\mathbf{x}, \mathbf{y}, \mathbf{s}) \in \mathcal{E} \cap \mathcal{C}(t)$*, and its first and second derivatives at* $(\mathbf{x}, \mathbf{y}, \mathbf{s})$ *be* $(\dot{\mathbf{x}}, \dot{\mathbf{y}}, \dot{\mathbf{s}})$ *and* $(\ddot{\mathbf{x}}, \ddot{\mathbf{y}}, \ddot{\mathbf{s}})$ *which are defined by (7) and (8). Then, the ellipse approximation of* (5) *is given by*

$$\mathbf{x}(\alpha) = \mathbf{x} - \dot{\mathbf{x}} \sin(\alpha) + \ddot{\mathbf{x}}(1 - \cos(\alpha)). \tag{9}$$

$$\mathbf{y}(\alpha) = \mathbf{y} - \dot{\mathbf{y}} \sin(\alpha) + \ddot{\mathbf{y}}(1 - \cos(\alpha)). \tag{10}$$

$$\mathbf{s}(\alpha) = \mathbf{s} - \dot{\mathbf{s}} \sin(\alpha) + \ddot{\mathbf{s}}(1 - \cos(\alpha)). \tag{11}$$

3 Infeasible predictor-corrector algorithm

We denote the duality measure by

$$\mu = \frac{\mathbf{x}^{\mathrm{T}}\mathbf{s}}{n}, \tag{12}$$

and define the set of neighborhood by

$$\mathcal{N}(\theta) := \{(\mathbf{x}, \mathbf{s}) \mid (\mathbf{x}, \mathbf{s}) > 0, \quad \|\mathbf{x} \circ \mathbf{s} - \mu\mathbf{e}\| \le \theta\mu\}. \tag{13}$$

The proposed algorithm searches an optimizer along the ellipse while staying inside $\mathcal{N}(\theta)$.

Algorithm 3.1
Data: $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$, $\theta \in (0, \frac{1}{2+\sqrt{2}}]$, $\epsilon > 0$, *initial point* $(\mathbf{x}^0, \mathbf{y}^0, \mathbf{s}^0) \in \mathcal{N}(\theta)$.
for *iteration* $k = 1, 2, \ldots$

Step 1: If

$$\mu_k \le \epsilon, \tag{14a}$$

$$\|\mathbf{r}_b^k\| = \|\mathbf{A}\mathbf{x}^k - \mathbf{b}\| \le \epsilon, \tag{14b}$$

$$\|\mathbf{r}_c^k\| = \|\mathbf{A}^{\mathrm{T}}\mathbf{y}^k + \mathbf{s}^k - \mathbf{c}\| \le \epsilon, \tag{14c}$$

$$(\mathbf{x}^k, \mathbf{s}^k) > 0. \tag{14d}$$

stop. Otherwise continue.

Step 2: Solve the linear systems of equations (7) and (8) to get $(\dot{\mathbf{x}}, \dot{\mathbf{y}}, \dot{\mathbf{s}})$ *and* $(\ddot{\mathbf{x}}, \ddot{\mathbf{y}}, \ddot{\mathbf{s}})$.

Step 3: Find the smallest positive $\bar{\alpha} \in (0, \pi/2]$ *such that for all* $\alpha \in (0, \bar{\alpha}]$, $(\mathbf{x}(\alpha), \mathbf{s}(\alpha)) > \mathbf{0}$ *and*

$$\|(\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha)) - (1 - \sin(\alpha))\mu_k \mathbf{e}\| \leq 2\theta(1 - \sin(\alpha))\mu_k. \tag{15}$$

Set (to simplify the notation, we use α *in stead of* $\bar{\alpha}$ *in the rest of the paper)*

$$(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) = (\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k) - (\dot{\mathbf{x}}, \dot{\mathbf{y}}, \dot{\mathbf{s}}) \sin(\alpha) + (\ddot{\mathbf{x}}, \ddot{\mathbf{y}}, \ddot{\mathbf{s}})(1 - \cos(\alpha)). \tag{16}$$

Step 4: Calculate $(\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s})$ *by solving*

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}^{\mathrm{T}} & \mathbf{I} \\ \mathbf{S}(\alpha) & \mathbf{0} & \mathbf{X}(\alpha) \end{bmatrix} \begin{bmatrix} \Delta\mathbf{x} \\ \Delta\mathbf{y} \\ \Delta\mathbf{s} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ (1 - \sin(\alpha))\mu_k \mathbf{e} - \mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) \end{bmatrix}. \tag{17}$$

Update

$$(\mathbf{x}^{k+1}, \mathbf{y}^{k+1}, \mathbf{s}^{k+1}) = (\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha)) + (\Delta\mathbf{x}, \Delta\mathbf{y}, \Delta\mathbf{s}) \tag{18}$$

and

$$\mu_{k+1} = \frac{{\mathbf{x}^{k+1}}^{\mathrm{T}} \mathbf{s}^{k+1}}{n}. \tag{19}$$

Step 5: Set $k + 1 \to k$. *Go back to Step 1.*

end (for) ■

In the rest of this section, we will show (1) $\mathbf{r}_b^k \to 0$, $\mathbf{r}_c^k \to 0$, and $\mu_k \to 0$; (2) there exist $\alpha \in (0, \pi/2]$ such that $(\mathbf{x}(\alpha), \mathbf{s}(\alpha)) > 0$ and (15) holds; (3) $(\mathbf{x}^k, \mathbf{s}^k) \in \mathcal{N}(\theta)$. It is easy to show that $\mathbf{r}_b^k$, $\mathbf{r}_c^k$, and μ_k decrease at the same rate in every iteration.

Lemma 3.1

$$\mathbf{r}_b^{k+1} = \mathbf{r}_b^k(1 - \sin(\alpha)), \quad \mathbf{r}_c^{k+1} = \mathbf{r}_c^k(1 - \sin(\alpha)), \quad \mu_{k+1} = \mu_k(1 - \sin(\alpha)). \tag{20}$$

Proof: Using (4), (18), (16), (8), and (7), we have

$$\begin{aligned} & \mathbf{r}_b^{k+1} - \mathbf{r}_b^k = \mathbf{A}(\mathbf{x}^{k+1} - \mathbf{x}^k) = \mathbf{A}(\mathbf{x}(\alpha) + \Delta\mathbf{x} - \mathbf{x}^k) \\ = \; & \mathbf{A}(\mathbf{x}^k - \dot{\mathbf{x}} \sin(\alpha) - \mathbf{x}^k) = -\mathbf{A}\dot{\mathbf{x}} \sin(\alpha) = -\mathbf{r}_b^k \sin(\alpha). \end{aligned}$$

This shows the first relation. The second relation follows a similar derivation. From (17), it holds that $(\Delta\mathbf{x})^{\mathrm{T}}\Delta\mathbf{s} = (\Delta\mathbf{x})^{\mathrm{T}}(-\mathbf{A}^{\mathrm{T}}\Delta\mathbf{y}) = -(\mathbf{A}\Delta\mathbf{x})^{\mathrm{T}}\Delta\mathbf{y} = 0$. Using (18), we have

$$\begin{aligned} & {\mathbf{x}^{k+1}}^{\mathrm{T}} \mathbf{s}^{k+1} = (\mathbf{x}(\alpha) + \Delta\mathbf{x})^{\mathrm{T}}(\mathbf{s}(\alpha) + \Delta\mathbf{s}) = \mathbf{x}(\alpha)^{\mathrm{T}}\mathbf{s}(\alpha) + \mathbf{x}(\alpha)^{\mathrm{T}}\Delta\mathbf{s} + \mathbf{s}(\alpha)^{\mathrm{T}}\Delta\mathbf{x} \\ = \; & \mathbf{x}(\alpha)^{\mathrm{T}}\mathbf{s}(\alpha) + (1 - \sin(\alpha))\mu_k n - \mathbf{x}(\alpha)^{\mathrm{T}}\mathbf{s}(\alpha) = (1 - \sin(\alpha))\mu_k n. \end{aligned}$$

Dividing both sides by n proves the last relation. ■

Clearly, if $\sin(\alpha) = 1$ ($\alpha = \frac{\pi}{2}$), we will find the optimal solution (allowing some $x_i = 0$ and/or $s_j = 0$) in one step, which is rarely the case. Therefore, from now on, we assume $\alpha \in (0, \frac{\pi}{2})$. We will use the following lemma of [16].

Lemma 3.2 *Let* $(\Delta\mathbf{x}, \Delta\mathbf{s})$ *be given by (17). Then*

$$\|\Delta\mathbf{x} \circ \Delta\mathbf{s}\| \le \frac{\sqrt{2}}{4}\|(\mathbf{X}(\alpha)\mathbf{S}(\alpha))^{-\frac{1}{2}}(\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) - \mu_{k+1}\mathbf{e})\|^2. \tag{21}$$

Theorem 3.1 *Let* $(\mathbf{x}(\alpha), \mathbf{y}(\alpha), \mathbf{s}(\alpha))$ *and* $(\mathbf{x}^k, \mathbf{s}^k) \in \mathcal{N}(\theta)$. *Then, for all* $k \ge 0$

(i) there is an $\alpha > 0$, *such that* $(\mathbf{x}(\alpha), \mathbf{s}(\alpha)) > 0$ *and (15) holds.*

(ii) if $\theta \le \frac{1}{2+\sqrt{2}}$, *then* $(\mathbf{x}^{k+1}, \mathbf{s}^{k+1}) \in \mathcal{N}(\theta)$ *for all the iterations.*

Proof: Using $1 - \cos(\alpha) \le 1 - \cos^2(\alpha) = \sin^2(\alpha)$ and $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k) \in \mathcal{N}(\theta)$, we have

$$\begin{aligned}
& \|\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) - (1 - \sin(\alpha))\mu_k\mathbf{e}\| \\
= \; & \|(\mathbf{x}^k \circ \mathbf{s}^k - \mu_k\mathbf{e})(1 - \sin(\alpha)) + (\ddot{\mathbf{x}} \circ \ddot{\mathbf{s}} - \dot{\mathbf{x}} \circ \dot{\mathbf{s}})(1 - \cos(\alpha))^2 \\
& -(\ddot{\mathbf{x}} \circ \dot{\mathbf{s}} + \dot{\mathbf{x}} \circ \ddot{\mathbf{s}}) \sin(\alpha)(1 - \cos(\alpha))\| \\
\le \; & \theta\mu_k(1 - \sin(\alpha)) + (\|\ddot{\mathbf{x}} \circ \ddot{\mathbf{s}}\| + \|\dot{\mathbf{x}} \circ \dot{\mathbf{s}}\|) \sin^4(\alpha) \\
& +(\|\ddot{\mathbf{x}} \circ \dot{\mathbf{s}}\| + \|\dot{\mathbf{x}} \circ \ddot{\mathbf{s}}\|) \sin^3(\alpha).
\end{aligned} \tag{22}$$

Clearly, if

$$q(\alpha) := \left(\left\|\ddot{\mathbf{x}} \circ \ddot{\mathbf{s}}\right\| + \left\|\dot{\mathbf{x}} \circ \dot{\mathbf{s}}\right\|\right) \sin^4(\alpha) + \left(\left\|\dot{\mathbf{x}} \circ \ddot{\mathbf{s}}\right\| + \left\|\ddot{\mathbf{x}} \circ \dot{\mathbf{s}}\right\|\right) \sin^3(\alpha) + \theta\mu_k \sin(\alpha) - \theta\mu_k \le 0, \tag{23}$$

then, (15) holds. Indeed, since $q(0) = -\theta\mu < 0$, by continuity, there exist $\alpha > 0$ such that (23) holds. This shows that (15) holds. From (15), we have

$$x_i(\alpha)s_i(\alpha) \ge (1 - 2\theta)(1 - \sin(\alpha))\mu_k > 0, \quad \forall\theta \in [0, 0.5) \text{ and } \forall\alpha \in [0, \pi/2).$$

This shows $(\mathbf{x}(\alpha), \mathbf{s}(\alpha)) > 0$. Therefore, we finish part (i). Furthermore, from Lemma 3.1, (15) is now equivalent to $\|\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) - \mu_{k+1}\mathbf{e}\| \le 2\theta\mu_{k+1}$. Using (18), (17), Lemmas 3.1 and 3.2, and part (i) of this theorem, we have

$$\begin{aligned}
& \|\mathbf{x}^{k+1} \circ \mathbf{s}^{k+1} - \mu_{k+1}\mathbf{e}\| \\
= \; & \|(\mathbf{x}(\alpha) + \Delta\mathbf{x}) \circ (\mathbf{s}(\alpha) + \Delta\mathbf{s}) - \mu_{k+1}\mathbf{e}\| \\
= \; & \|\Delta\mathbf{x} \circ \Delta\mathbf{s}\| \le \frac{\sqrt{2}}{4}\|(\mathbf{X}(\alpha)\mathbf{S}(\alpha))^{-\frac{1}{2}}(\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) - \mu_{k+1}\mathbf{e})\|^2 \\
\le \; & \frac{\sqrt{2}}{4}\frac{\|\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) - \mu_{k+1}\mathbf{e}\|^2}{\min_i x_i(\alpha)s_i(\alpha)} \\
\le \; & \frac{\sqrt{2}(2\theta)^2\mu_{k+1}^2}{4(1 - 2\theta)\mu_{k+1}} \\
\le \; & \frac{\sqrt{2}\theta^2}{(1 - 2\theta)}\mu_{k+1}.
\end{aligned} \tag{24}$$

It is easy to check that for $\theta \le \frac{1}{2+\sqrt{2}} \approx 0.29289$, $\frac{\sqrt{2}\theta^2}{(1-2\theta)} \le \theta$ holds, therefore, for $\theta \le \frac{1}{2+\sqrt{2}}$, we have

$$\|\mathbf{x}^{k+1} \circ \mathbf{s}^{k+1} - \mu_{k+1}\mathbf{e}\| \le \theta\mu_{k+1}.$$

We now show that $(\mathbf{x}^{k+1}, \mathbf{s}^{k+1}) > 0$. Let $\mathbf{x}^{k+1}(t) = \mathbf{x}(\alpha) + t\Delta\mathbf{x}$ and $\mathbf{s}^{k+1}(t) = \mathbf{s}(\alpha) + t\Delta\mathbf{s}$. Then, $\mathbf{x}^{k+1}(0) = \mathbf{x}(\alpha)$ and $\mathbf{x}^{k+1}(1) = \mathbf{x}^{k+1}$. Since

$$\begin{aligned} & \mathbf{x}^{k+1}(t) \circ \mathbf{s}^{k+1}(t) = (\mathbf{x}(\alpha) + t\Delta\mathbf{x}) \circ (\mathbf{s}(\alpha) + t\Delta\mathbf{s}) \\ = \; & \mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) + t(\mathbf{x}(\alpha) \circ \Delta\mathbf{s} + \mathbf{s}(\alpha) \circ \Delta\mathbf{x}) + t^2\Delta\mathbf{x} \circ \Delta\mathbf{s}, \end{aligned}$$

using (17), (15), (24), and the assumption that $\theta \le \frac{1}{2+\sqrt{2}}$, we have

$$\begin{aligned} & \|\mathbf{x}^{k+1}(t) \circ \mathbf{s}^{k+1}(t) - \mu_{k+1}\mathbf{e}\| \\ = \; & \|(1-t)(\mathbf{x}(\alpha) \circ \mathbf{s}(\alpha) - \mu_{k+1}\mathbf{e}) + t^2\Delta\mathbf{x} \circ \Delta\mathbf{s}\| \\ \le \; & 2(1-t)\theta\mu_{k+1} + t^2\frac{\sqrt{2}\theta^2}{1-2\theta}\mu_{k+1} \\ \le \; & (2(1-t) + t^2)\theta\mu_{k+1} := f(t)\theta\mu_{k+1}. \end{aligned} \tag{25}$$

The function $f(t)$ is a monotonical decreasing function of $t \in [0,1]$, and $f(0) = 2$. This proves $\|\mathbf{x}^{k+1}(t) \circ \mathbf{s}^{k+1}(t) - \mu_{k+1}\mathbf{e}\| \le 2\theta\mu_{k+1}$. Therefore, $x_i^{k+1}(t)s_i^{k+1}(t) \ge (1-2\theta)\mu_{k+1} > 0$ for all $t \in [0,1]$, which means $(\mathbf{x}^{k+1}\mathbf{s}^{k+1}) > 0$. This finishes the proof of part (ii). ∎

This theorem indicates that the proposed algorithm is well-defined.

4 Polynomiality

The analysis follows similar ideas in many existing literatures, such as [16, 22]. Let the initial point be selected to satisfy

$$(\mathbf{x}^0, \mathbf{s}^0) \in \mathcal{N}(\theta), \;\; \mathbf{x}^* \le \rho\mathbf{x}^0, \;\; \mathbf{s}^* \le \rho\mathbf{s}^0, \;\; (\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*) \in \mathcal{S}, \tag{26}$$

where $\rho \ge 1$. Let ω^f and ω^o be the quality of the initial point which are the "distances" from feasibility and optimility given by

$$\omega^f = \min_{\mathbf{x},\mathbf{y},\mathbf{s}} \{\max\{\|(\mathbf{X}^0)^{-1}(\mathbf{x}-\mathbf{x}^0)\|_\infty, \|(\mathbf{S}^0)^{-1}(\mathbf{s}-\mathbf{s}^0)\|_\infty\} \;|\; \mathbf{A}\mathbf{x} = \mathbf{b}, \;\; \mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{s} = \mathbf{c}\}. \tag{27}$$

and

$$\omega^0 = \min_{\mathbf{x}^*,\mathbf{y}^*,\mathbf{s}^*} \{\max\{\frac{\mathbf{x}^{*\mathrm{T}}\mathbf{s}^0}{\mathbf{x}^{0\mathrm{T}}\mathbf{s}^0}, \frac{\mathbf{s}^{*\mathrm{T}}\mathbf{x}^0}{\mathbf{x}^{0\mathrm{T}}\mathbf{s}^0}, 1\} \;|\; (\mathbf{x}^*, \mathbf{y}^*, \mathbf{s}^*) \in \mathcal{S}\}. \tag{28}$$

Let ω_p^r and ω_d^r be the "ratios" of the feasibility and the total complementarity defined by

$$\omega_p^r = \frac{\|\mathbf{A}\mathbf{x}^0 - \mathbf{b}\|}{\mathbf{x}^{0\mathrm{T}}\mathbf{s}^0}, \tag{29a}$$

$$\omega_d^r = \frac{\|\mathbf{A}^{\mathrm{T}}\mathbf{y}^0 + \mathbf{s}^0 - \mathbf{c}\|}{\mathbf{x}^{0\mathrm{T}}\mathbf{s}^0}. \tag{29b}$$

In view of Lemma 3.1, we have that

$$\|\mathbf{A}\mathbf{x}^k - \mathbf{b}\| = \omega_p^r {\mathbf{x}^k}^{\mathrm{T}} \mathbf{s}^k, \tag{30a}$$

$$\|\mathbf{A}^{\mathrm{T}}\mathbf{y}^k + \mathbf{s}^k - \mathbf{c}\| = \omega_d^r {\mathbf{x}^k}^{\mathrm{T}} \mathbf{s}^k. \tag{30b}$$

Invoking Lemma 3.3 of [22] for $\lambda_p = \lambda_d = \xi = 1$ and (7), we have the following two Lemmas [16].

Lemma 4.1 *Let $(\dot{\mathbf{x}}, \dot{\mathbf{s}})$ be defined by (7), and $\mathbf{D}^k = (\mathbf{X}^k)^{\frac{1}{2}}(\mathbf{S}^k)^{-\frac{1}{2}}$. Then*

$$\max\{\|(\mathbf{D}^k)^{-1}\dot{\mathbf{x}}\|, \|(\mathbf{D}^k)\dot{\mathbf{s}}\|\} \le \|(\mathbf{x}^k \circ \mathbf{s}^k)^{\frac{1}{2}}\| + \omega^f(1+2\omega^o)\frac{(\mathbf{x}^k)^{\mathrm{T}}\mathbf{s}^k}{\min_i(x_i^k s_i^k)^{\frac{1}{2}}}. \tag{31}$$

Lemma 4.2 *Let $(\mathbf{x}^0, \mathbf{s}^0)$ be defined by (26). Then*

$$\omega^f \le \rho, \quad \omega^o \le \rho. \tag{32}$$

This leads to the following lemma.

Lemma 4.3 *Let $(\dot{\mathbf{x}}, \dot{\mathbf{s}})$ be defined by (7). Then, there exists a positive constant C_0, independent of n, such that*

$$\max\{\|(\mathbf{D}^k)^{-1}\dot{\mathbf{x}}\|, \|(\mathbf{D}^k)\dot{\mathbf{s}}\|\} \le C_0\sqrt{n(\mathbf{x}^k)^{\mathrm{T}}\mathbf{s}^k}. \tag{33}$$

Proof: First, it is easy to see

$$\|(\mathbf{x}^k \circ \mathbf{s}^k)^{\frac{1}{2}}\| = \sqrt{\sum_i x_i^k s_i^k} = \sqrt{(\mathbf{x}^k)^{\mathrm{T}}\mathbf{s}^k}. \tag{34}$$

Since $(\mathbf{x}^k, \mathbf{s}^k) \in \mathcal{N}(\theta)$, we have $\min_i(x_i^k s_i^k) \ge (1-\theta)\mu_k = (1-\theta)\frac{(\mathbf{x}^k)^{\mathrm{T}}\mathbf{s}^k}{n}$. Therefore,

$$\frac{(\mathbf{x}^k)^{\mathrm{T}}\mathbf{s}^k}{\min_i(x_i^k s_i^k)^{\frac{1}{2}}} \le \sqrt{\frac{n(\mathbf{x}^k)^{\mathrm{T}}\mathbf{s}^k}{(1-\theta)}}. \tag{35}$$

Substituting (34) and (35) into (31) and using Lemma 4.2 prove (33) with $C_0 = 1 + \frac{\rho(1+2\rho)}{\sqrt{(1-\theta)}} \ge 1 + \frac{\omega^f(1+2\omega^o)}{\sqrt{(1-\theta)}}$. ∎

From Lemma 4.3, we can establish several useful inequalities. The following simple facts will be used several times. Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors, then

$$\|\mathbf{u} \circ \mathbf{v}\|^2 = \sum_i (u_i v_i)^2 \leq \left(\sum_i u_i^2\right)\left(\sum_i v_i^2\right) = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2. \tag{36}$$

If $\mathbf{u}$ and $\mathbf{v}$ satisfy $\mathbf{u}^{\mathrm{T}}\mathbf{v} = 0$, then,

$$\max\{\|\mathbf{u}\|^2, \|\mathbf{v}\|^2\} \leq \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \|\mathbf{u} + \mathbf{v}\|^2, \tag{37}$$

and (see [23, Lemma 5.3])

$$\|\mathbf{u} \circ \mathbf{v}\| \leq 2^{-\frac{3}{2}} \|\mathbf{u} + \mathbf{v}\|^2. \tag{38}$$

Lemma 4.4 *Let $(\dot{\mathbf{x}}, \dot{\mathbf{s}})$ and $(\ddot{\mathbf{x}}, \ddot{\mathbf{s}})$ be defined by (7) and (8), respectively. Then, there exist positive constants C_1, C_2, C_3, and C_4, independent of n, such that*

$$\|\dot{\mathbf{x}} \circ \dot{\mathbf{s}}\| \leq C_1 n^2 \mu_k, \tag{39}$$

$$\|\ddot{\mathbf{x}} \circ \ddot{\mathbf{s}}\| \leq C_2 n^4 \mu_k, \tag{40}$$

$$\max\{\|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}}\|, \|(\mathbf{D}^k)\ddot{\mathbf{s}}\|\} \leq C_3 n^2 \sqrt{\mu_k}, \tag{41}$$

$$\max\{\|\ddot{\mathbf{x}} \circ \dot{\mathbf{s}}\|, \|\dot{\mathbf{x}} \circ \ddot{\mathbf{s}}\|\} \leq C_4 n^3 \mu_k \tag{42}$$

Proof: First, using (36) and Lemma 4.3, we have

$$\|\dot{\mathbf{x}} \circ \dot{\mathbf{s}}\| = \|(\mathbf{D}^k)^{-1}\dot{\mathbf{x}} \circ (\mathbf{D}^k)\dot{\mathbf{s}}\| \leq \|(\mathbf{D}^k)^{-1}\dot{\mathbf{x}}\| \|(\mathbf{D}^k)\dot{\mathbf{s}}\| \leq C_0^2 n (\mathbf{x}^k)^{\mathrm{T}} \mathbf{s}^k := C_1 n^2 \mu_k. \tag{43}$$

Second, using (38), (8), (39), and (34), we have

$$\begin{aligned}
&\|\ddot{\mathbf{x}} \circ \ddot{\mathbf{s}}\| = \|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}} \circ (\mathbf{D}^k)\ddot{\mathbf{s}}\| \leq 2^{-\frac{3}{2}} \|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}} + (\mathbf{D}^k)\ddot{\mathbf{s}}\|^2 \\
\leq\ & 2^{-\frac{3}{2}} \left\| -2(\mathbf{XS})^{-\frac{1}{2}} (\dot{\mathbf{x}} \circ \dot{\mathbf{s}}) \right\|^2 \\
=\ & 2^{\frac{1}{2}} \sum_{i=1}^{n} \left(\frac{\dot{x}_i \dot{s}_i}{\sqrt{x_i}\sqrt{s_i}} \right)^2 = 2^{\frac{1}{2}} \sum_{i=1}^{n} \frac{(\dot{x}_i \dot{s}_i)^2}{x_i s_i} \\
\leq\ & 2^{\frac{1}{2}} \frac{\sum_{i=1}^{n} (\dot{x}_i \dot{s}_i)^2}{\min_{i=1,\ldots,n} x_i s_i} \\
\leq\ & 2^{\frac{1}{2}} \frac{||\dot{x} \circ \dot{s}||^2}{(1-\theta)\mu_k} \leq 2^{\frac{1}{2}} \frac{C_1^2 n^4 \mu_k^2}{(1-\theta)\mu_k} \\
=\ & 2^{\frac{1}{2}} \frac{C_1^2 n^4 \mu_k}{1-\theta} := C_2 n^4 \mu_k.
\end{aligned} \tag{44}$$

Third, using (37), (8), and (39), we have

$$\begin{aligned}
&\max\{\|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}}\|^2, \|(\mathbf{D}^k)\ddot{\mathbf{s}}\|^2\} \leq \|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}} + (\mathbf{D}^k)\ddot{\mathbf{s}}\|^2 \\
=\ & \left\| -2(\mathbf{XS})^{-\frac{1}{2}} (\dot{\mathbf{x}} \circ \dot{\mathbf{s}}) \right\|^2 \leq \frac{4C_1^2 n^4 \mu_k}{1-\theta} := C_3^2 n^4 \mu_k.
\end{aligned} \tag{45}$$

Taking square root on both sides proves (41). Finally, using (36), (41), and Lemma 4.3, we have

$$\begin{aligned} \|\ddot{\mathbf{x}} \circ \dot{\mathbf{s}}\| &= \|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}} \circ (\mathbf{D}^k)\dot{\mathbf{s}}\| \leq \|(\mathbf{D}^k)^{-1}\ddot{\mathbf{x}}\| \|(\mathbf{D}^k)\dot{\mathbf{s}}\| \\ &\leq (C_3 n^2 \sqrt{\mu_k})(C_0 n \sqrt{\mu_k}) := C_4 n^3 \mu_k. \end{aligned} \tag{46}$$

Similarly, we can show

$$\|\dot{\mathbf{x}} \circ \ddot{\mathbf{s}}\| \leq C_4 n^3 \mu_k. \tag{47}$$

This finishes the proof. ∎

Now we are ready to estimate a conservative bound for $\sin(\alpha)$.

Lemma 4.5 *Let* $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k)$ *be generated by Algorithm 3.1. Then,* $\sin(\alpha)$ *obtained in Step 3 satisfies the following inequality.*

$$\sin(\alpha) \geq \frac{\theta}{2Cn}, \tag{48}$$

where $C = \max\{1, C_4^{\frac{1}{3}}, (C_1 + C_2)^{\frac{1}{4}}\}$.

Proof: Let $\sin(\alpha) = \frac{\theta}{2Cn}$. In view of (23) and Lemma 4.4, we have

$$\begin{aligned} q(\alpha) &\leq \mu_k((C_1 + C_2)n^4 \sin^4(\alpha) + 2C_4 n^3 \sin^3(\alpha) + \theta \sin(\alpha) - \theta) := \mu_k p(\alpha) \\ &\leq \mu_k \left(\frac{(C_1 + C_2)\theta^4}{16C^4} + \frac{2C_4\theta^3}{8C^3} + \frac{\theta^2}{2Cn} - \theta \right) \\ &\leq \mu_k \left(\frac{\theta^4}{16} + \frac{\theta^3}{4} + \frac{\theta^2}{2} - \theta \right) \leq 0. \end{aligned}$$

Since $p(\alpha)$ is a monotonic function of $\sin(\alpha)$, for all $\sin(\alpha) \leq \frac{\theta}{2Cn}$, the above inequalities hold (the last inequality holds because of $\theta \leq 1$). Therefore, for all $\sin(\alpha) \leq \frac{\theta}{2Cn}$, the inequality (15) holds. This finishes the proof. ∎

Remark 4.1 *It is worthwhile to point out that the constant* C *depends on* C_0 *which depends on* ρ*, but* ρ *is an unknown before we find the solution. Also, we can always find a better steplength* $\sin(\alpha)$ *by solving the quartic* $q(\alpha) = 0$ *and the calculation of the roots for a quartic polynomial is deterministic, negligible, and independent to* n *[24, 25].*

Following the standard argument developed in [23], we have the main theorem.

Theorem 4.1 *Let* $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k)$ *be generated by Algorithm 3.1 with an initial point given by (26). For any* $\epsilon > 0$*, the algorithm will terminate with* $(\mathbf{x}^k, \mathbf{y}^k, \mathbf{s}^k)$ *satisfying (14) in at most* $\mathcal{O}(nL)$ *iterations, where*

$$L = \max\{\ln((\mathbf{x}^0)^{\mathrm{T}}\mathbf{s}^0/\epsilon), \ln(\|\mathbf{r}_b^0\|/\epsilon), \ln(\|\mathbf{r}_c^0\|/\epsilon)\}.$$

Proof: In view of Lemma 3.1, $\mathbf{r}_b^k$, $\mathbf{r}_c^k$, and μ_k decrease at the same rate $(1 - \sin(\alpha))$ in every iteration. Using the Lemma 4.5 and [23, Theorem 3.2] proves the claim. ∎

5 Conclusions

We proposed an infeasible-interior-point algorithm that searches the optimizer along an ellipse that approximates the central-path. We showed that the proposed algorithm is polynomial and that the polynomial bound is at least as good as the best existing bound for infeasible-interior-point algorithms for linear programming.

References

[1] V. Klee and G. Minty, How good is the simplex algorithm? In: Shisha, O., (eds.) Inequalities, Vol. III, pp. 159-175, Academic Press, (1972).

[2] L. Khachiyan, A polynomial algorithm in linear programming, Doklady Akademiia Nauk SSSR, 224, 1093-1096 (1979).

[3] R. Bland, D. Goldfarb, and M. Todd, The ellipsoid method: A survey, Operations Research, 29, 1039-1091 (1981).

[4] N. Karmarkar, A new polynomial-time algorithm for linear programming, Combinatorica, 4, 375-395 (1984).

[5] J. Renegar, A polynomial-time algorithm based on Newton's method, for linear programming, Mathematical Programming, 40, 59-93 (1988).

[6] M. Kojima, S. Mizuno, and A. Yoshise, A polynomial-time algorithm for a class of linear complementarity problem, Mathematical Programming, 44, 1-26 (1989).

[7] M. Kojima, S. Mizuno, and A. Yoshise, A primal-dual interior point algorithm for linear programming, in Progress in Mathematical Programming: Interior-point and Related Methods, N. Megiddo, ed., Springer-Verlag, New York, 1989.

[8] C. Gonzaga, Polynomial affine algorithms for linear programming, Mathematical Programming, 49, 6-21 (1990).

[9] Y. Ye, An $\mathcal{O}(n^3L)$ potential reduction algorithm for linear programming, Mathematical Programming, 50, 239-258 (1991).

[10] S. Mizuno, M. Todd, and Y. Ye, On adaptive step primal-dual interior-point algorithms for linear programming, Mathematics of Operations Research, 18, 964-981 (1993).

[11] S. Mehrotra, On the implementation of a primal-dual interior point method, SIAM Journal on Optimization, 2, 575-601 (1992).

[12] I. Lustig, R. Marsten, and D. Shannon, On implementing Mehrotra's predictor-corrector interior-point method for linear programming, SIAM Journal on Optimization, 2, 432-449 (1992).

[13] R. Monteiro, I. Adler, and M. Resende, A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension, Mathematics of Operations Research, 15, 191-214 (1990).

[14] Y. Zhang, On the convergence of a class of infeasible-interior-point methods for the horizontal linear Complementarity problem, SIAM Journal on Optimization, 4, 208-227 (1994).

[15] S. Mizuno, Polynomiality of the Kojima-Megiddo-Mizuno infeasible interior point algorithm for linear programming, Mathematical Programming, 67, 109-119 (1994)

[16] J. Miao, Two infeasible interior-point predictor-corrector algorithms for linear programming, SIAM Journal on Optimization, 6, 587-599 (1996).

[17] Y. Yang, A Polynomial Arc-Search Interior-Point Algorithm for Linear Programming, Journal of Optimization Theory and Applications, 158, 859-873 (2013).

[18] Y. Yang, A polynomial arc-search interior-point algorithm for convex quadratic programming, European Journal of Operational Research, 215, 25-38 (2011).

[19] X. Yang, Y. Zhang, and H. Liu, A wide neighborhood infeasible-interior-point method with arc-search for linear programming, Journal of Applied Mathematics and Computing, 1-17 (2015).

[20] M. P. Do Carmo, Differential Geometry of Curves and Surfaces, Prentice-Hall, New Jersey, (1976).

[21] S. Mizuno, A new polynomial time method for a linear complemenetarity problem, Mathematical Programming 56, 31-43 (1992).

[22] M. Kojima, Basic lemmas in polynomial-time infeasible-interior-point methods for linear programming, Annals of Operations Research, 62, 1-28 (1996).

[23] S. Wright, Primal-Dual Interior-Point Methods, SIAM, Philadelphia, (1997).

[24] S.L. Shmakov, A universal method of solving quartic equations, International Journal of Pure and Applied Mathematics, 71 (2), 251-259 (2011).

[25] Y. Yang and Z. Zhou, An analytic solution to Wahba's problem, Aerospace Science and Technology, 30, 46-49 (2013)